



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Thompson, Gregory A
Knauf, Vic C
- (ii) TITLE OF INVENTION: Plant Desaturases-Compositions
and Uses
- (iii) NUMBER OF SEQUENCES: 43
- (iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Calgene, Inc.
- (B) STREET: 1920 Fifth Street
- (C) CITY: Davis
- (D) STATE: California
- (E) COUNTRY: USA
- (F) ZIP: 95616

RECEIVED
APR 29 1992
GROUP 180

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB storage
- (B) COMPUTER: Apple Macintosh
- (C) OPERATING SYSTEM: Macintosh 6.0.7
- (D) SOFTWARE: MicrosoftWord 4.0

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 07/762,762
- (B) FILING DATE: 16-SEPT-1991
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: PCT/US91/01746
- (B) FILING DATE: 14-MAR-1991
- (C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

T1160

(A) APPLICATION NUMBER: 07/615,784

(B) FILING DATE: 14-NOV-1990

(A) APPLICATION NUMBER: 07/567,373

(B) FILING DATE: 13-AUG-1990

(A) APPLICATION NUMBER: 07/494,106

(B) FILING DATE: 16-MAR-1990

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Lassen, Elizabeth

(B) REGISTRATION NUMBER: 31,845

(A) NAME: Donna E. Scherer

(B) REGISTRATION NUMBER: 34,719

(C) REFERENCE/DOCKET NUMBER: CGNE 69-4

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (916) 753-6313

(B) TELEFAX: (916) 753-1510

(C) TELEX: 350370 CGNE

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala	Ser	Thr	Leu	Gly	Ser	Ser	Thr	Pro	Lys	Val	Asp	Asn	Ala	Lys	Lys
1				5					10					15	
Pro	Phe	Gln	Pro	Pro	Arg	Glu	Val	His	Val	Gln	Val	Thr	His	Xaa	Met
			20					25						30	

Pro Pro Gln Lys Ile Glu Ile Phe Lys Ser Ile Glu Gly Xaa Ala Glu
 35 40 45
 Gln Asn Ile Leu Val Xaa Leu Lys Pro Val Glu Lys Cys Trp Gln
 50 55 60

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asp Phe Leu Pro Asp Pro Ala Xaa Glu Gly Phe Asp Glu Gln Val Lys
 1 5 10 15
 Glu Leu Arg Ala Arg Ala Lys Glu Ile Pro Asp Asp Tyr Phe Val Val
 20 25 30
 Leu Val Gly Asp Met Ile Thr Glu Glu Ala Leu Pro Thr Tyr Gln Thr
 35 40 45
 Met Leu Asn Thr Leu Asp Gly Val
 50 55

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Asp Glu Thr Gly Ala Ser Leu Thr Pro Trp Ala Val Trp Thr
 1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asp Leu Leu His Thr Tyr Leu Tyr Leu Ser Gly Arg Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Met Arg Gln Ile Gln Lys Thr Ile Gln Tyr Leu Ile
1 5 10

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Thr Glu Asn Ser Pro Tyr Leu Gly Phe Ile Tyr Thr Ser Phe Gln Glu
1 5 10 15

Arg

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp Val Xaa Leu Ala Gln Ile Xaa Gly Thr Ile Ala Ser Asp Glu Lys
1 5 10 15

Arg His Glu Thr Ala Tyr Thr Lys Ile Val Glu Lys Leu Phe Glu Ile
20 25 30

Asp Pro Asp Gly Thr Val Leu Ala Phe Ala Asp Met Met Arg Lys Lys
35 40 45

Ile Xaa Met Pro Ala His Leu Met Tyr
50 55

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Asp Asn Leu Phe
1

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Xaa	Xaa	Phe	Xaa	Ala	Val	Xaa	Gln	Arg	Leu	Xaa	Val	Tyr	Thr	Ala	Lys
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asp	Tyr	Ala	Asp	Ile	Leu	Glu	Phe	Leu	Val	Gly	Arg	Trp	Lys
1				5				10					

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Val Ala Asp Leu Thr Gly Leu Ser Gly Glu Gly Arg Lys Ala Xaa Asp
1 5 10 15
Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg Arg Leu Glu Glu Arg Ala
20 25 30
Gln Gly Arg Ala Lys Glu Gly Pro Val Val Pro Phe Ser Trp Ile Phe
35 40 45
Asp Arg Gln Val Lys Leu
50

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1533 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCTCACTTGT GTGGTGGAGG AGAAAAACAG AACTCACAAA AAGCTTTGCG ACTGCCAAGA 60
ACAACAACAA CAACAAGATC AAGAAGAAGA AGAAGAAGAT CAAAAATGGC TCTTCGAATC 120
ACTCCAGTGA CCTTGCAATC GGAGAGATAT CGTTCGTTTT CGTTTCCTAA GAAGGCTAAT 180
CTCAGATCTC CCAAATTCGC CATGGCCTCC ACCCTCGGAT CATCCACACC GAAGGTTGAC 240
AATGCCAAGA AGCCTTTTCA ACCTCCACGA GAGGTTCATG TTCAGGTGAC GCACTCCATG 300
CCACCACAGA AGATAGAGAT TTTCAAATCC ATCGAGGGTT GGGCTGAGCA GAACATATTG 360
GTTACCTTAA AGCCAGTGGA GAAATGTTGG CAAGCACAGG ATTTCTTGCC GGACCCTGCA 420
TCTGAAGGAT TTGATGAACA AGTCAAGGAA CTAAGGGCAA GAGCAAAGGA GATTCCTGAT 480
GATTACTTTG TTGTTTTGGT TGGAGATATG ATTACAGAGG AAGCCCTACC TACTTACCAA 540
ACAATGCTTA ATACCCTAGA TGGTGTACGT GATGAGACTG GGGCTAGCCT TACGCCTTGG 600
GCTGTCTGGA CTAGGGCTTG GACAGCTGAA GAGAACAGGC ATGGCGATCT TCTCCACACC 660

TATCTCTACC TTTCTGGGCG GGTAGACATG AGGCAGATAC AGAAGACAAT TCAGTATCTC 720
ATTGGGTCAG GAATGGATCC TCGTACCGAA AACAGCCCCT ACCTTGGGTT CATCTACACA 780
TCGTTTCAAG AGCGTGCCAC ATTTGTTTCT CACGGAAACA CCGCCAGGCA TGCAAAGGAT 840
CATGGGGACG TGAAACTGGC GCAAATTTGT GGTACAATCG CGTCTGACGA AAAGCGTCAC 900
GAGACCGCTT ATACAAAGAT AGTCGAAAAG CTATTCGAGA TCGATCCTGA TGGCACCGTT 960
CTTGCTTTTG CCGACATGAT GAGGAAAAAG ATCTCGATGC CCGCACACTT GATGTACGAT 1020
GGGCGTGATG ACAACCTCTT CGAACATTTT TCGGCGGTTG CCCAAAGACT CGGCGTCTAC 1080
ACCGCCAAAG ACTACGCCGA CATACTGGAA TTTCTGGTCG GGCGGTGGAA AGTGGCGGAT 1140
TTGACCGGCC TATCTGGTGA AGGGCGTAAA GCGCAAGATT ATGTTTGCGG GTTGCCACCA 1200
AGAATCAGAA GGCTGGAGGA GAGAGCTCAA GGGCGAGCAA AGGAAGGACC TGTTGTTCCA 1260
TTCAGCTGGA TTTTCGATAG ACAGGTGAAG CTGTGAAGAA AAAAAAACG AGCAGTGAGT 1320
TCGGTTTCTG TTGGCTTATT GGGTAGAGGT TAAAACCTAT TTTAGATGTC TGTTTCGTGT 1380
AATGTGGTTT TTTTCTTCT AATCTTGAAT CTGGTATTGT GTCGTTGAGT TCGCGTGTGT 1440
GTAAACTTGT GTGGCTGTGG ACATATTATA GAACTCGTTA TGCCAATTTT GATGACGGTG 1500
GTTATCGTCT CCCCTGGTGT TTTTTTATTG TTT 1533

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 396 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Ala Leu Arg Ile Thr Pro Val Thr Leu Gln Ser Glu Arg Tyr Arg
-30 -25 -20
Ser Phe Ser Phe Pro Lys Lys Ala Asn Leu Arg Ser Pro Lys Phe Ala
-15 -10 -5

Met	Ala	Ser	Thr	Leu	Gly	Ser	Ser	Thr	Pro	Lys	Val	Asp	Asn	Ala	Lys
1					5					10					15
Lys	Pro	Phe	Gln	Pro	Pro	Arg	Glu	Val	His	Val	Gln	Val	Thr	His	Ser
				20					25					30	
Met	Pro	Pro	Gln	Lys	Ile	Glu	Ile	Phe	Lys	Ser	Ile	Glu	Gly	Trp	Ala
			35					40					45		
Glu	Gln	Asn	Ile	Leu	Val	His	Leu	Lys	Pro	Val	Glu	Lys	Cys	Trp	Gln
		50					55					60			
Ala	Gln	Asp	Phe	Leu	Pro	Asp	Pro	Ala	Ser	Glu	Gly	Phe	Asp	Glu	Gln
65						70					75				
Val	Lys	Glu	Leu	Arg	Ala	Arg	Ala	Lys	Glu	Ile	Pro	Asp	Asp	Tyr	Phe
80				85						90					95
Val	Val	Leu	Val	Gly	Asp	Met	Ile	Thr	Glu	Glu	Ala	Leu	Pro	Thr	Tyr
				100					105					110	
Gln	Thr	Met	Leu	Asn	Thr	Leu	Asp	Gly	Val	Arg	Asp	Glu	Thr	Gly	Ala
			115					120					125		
Ser	Leu	Thr	Pro	Trp	Ala	Val	Trp	Thr	Arg	Ala	Trp	Thr	Ala	Glu	Glu
		130					135					140			
Asn	Arg	His	Gly	Asp	Leu	Leu	His	Thr	Tyr	Leu	Tyr	Leu	Ser	Gly	Arg
	145					150					155				
Val	Asp	Met	Arg	Gln	Ile	Gln	Lys	Thr	Ile	Gln	Tyr	Leu	Ile	Gly	Ser
160					165					170					175
Gly	Met	Asp	Pro	Arg	Thr	Glu	Asn	Ser	Pro	Tyr	Leu	Gly	Phe	Ile	Tyr
				180					185					190	
Thr	Ser	Phe	Gln	Glu	Arg	Ala	Thr	Phe	Val	Ser	His	Gly	Asn	Thr	Ala
			195					200					205		
Arg	His	Ala	Lys	Asp	His	Gly	Asp	Val	Lys	Leu	Ala	Gln	Ile	Cys	Gly
		210					215					220			
Thr	Ile	Ala	Ser	Asp	Glu	Lys	Arg	His	Glu	Thr	Ala	Tyr	Thr	Lys	Ile
	225					230					235				
Val	Glu	Lys	Leu	Phe	Glu	Ile	Asp	Pro	Asp	Gly	Thr	Val	Leu	Ala	Phe
240					245					250					255
Ala	Asp	Met	Met	Arg	Lys	Lys	Ile	Ser	Met	Pro	Ala	His	Leu	Met	Tyr
				260					265					270	
Asp	Gly	Arg	Asp	Asp	Asn	Leu	Phe	Glu	His	Phe	Ser	Ala	Val	Ala	Gln
			275					280					285		

Arg Leu Gly Val Tyr Thr Ala Lys Asp Tyr Ala Asp Ile Leu Glu Phe
290 295 300

Leu Val Gly Arg Trp Lys Val Ala Asp Leu Thr Gly Leu Ser Gly Glu
305 310 315

Gly Arg Lys Ala Gln Asp Tyr Val Cys Gly Leu Pro Pro Arg Ile Arg
320 325 330 335

Arg Leu Glu Glu Arg Ala Gln Gly Arg Ala Lys Glu Gly Pro Val Val
340 345 350

Pro Phe Ser Trp Ile Phe Asp Arg Gln Val Lys Leu
355 360

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 225 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAAAGAAAAA GGTAAGAAAA AAAACA ATG GCT CTC AAG CTC AAT CCT TTC CTT 53
MET Ala Leu Lys Leu Asn Pro Phe Leu
1 5

TCT CAA ACC CAA AAG TTA CCT TCT TTC GCT CTT CCA CCA ATG GCC AGT 101
Ser Gln Thr Gln Lys Leu Pro Ser Phe Ala Leu Pro Pro MET Ala Ser
10 15 20 25

ACC AGA TCT CCT AAG TTC TAC ATG GCC TCT ACC CTC AAG TCT GGT TCT 149
Thr Arg Ser Pro Lys Phe Tyr MET Ala Ser Thr Leu Lys Ser Gly Ser
30 35 40

AAG GAA GTT GAG AAT CTC AAG AAG CCT TTC ATG CCT CCT CGG GAG GTA 197
Lys Glu Val Glu Asn Leu Lys Lys Pro Phe MET Pro Pro Arg Glu Val
45 50 55

CAT GTT CAG GTT ACC CAT TCT ATT GCC A 225
His Val Gln Val Thr His Ser Ile Ala
60 65

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1668 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
AAAAGAAAAA GGTAAGAAAA AAAACAATGG CTCTCAAGCT CAATCCTTTC CTTTCTCAAA 60
CCCAAAAGTT ACCTTCTTTC GCTCTTCCAC CAATGGCCAG TACCAGATCT CCTAAGTTCT 120
ACATGGCCTC TACCCTCAAG TCTGGTTCTA AGGAAGTTGA GAATCTCAAG AAGCCTTTCA 180
TGCCTCCTCG GGAGGTACAT GTTCAGGTTA CCCATTCTAT GCCACCCCAA AAGATTGAGA 240
TCTTTAAATC CCTAGACAAT TGGGCTGAGG AGAACATTCT GGTTCATCTG AAGCCAGTTG 300
AGAAATGTTG GCAACCGCAG GATTTTTTGC CAGATCCCGC CTCTGATGGA TTTGATGAGC 360
AAGTCAGGGA ACTCAGGGAG AGAGCAAAGG AGATTCCTGA TGATTATTTT GTTGTTTTTG 420
TTGGAGACAT GATAACGGAA GAAGCCCTTC CCACTTATCA AACAATGCTG AATACCTTGG 480
ATGGAGTTCG GGATGAAACA GGTGCAAGTC CTACTTCTTG GGCAATTTGG ACAAGGGCAT 540
GGACTGCGGA AGAGAATAGA CATGGTGACC TCCTCAATAA GTATCTCTAC CTATCTGGAC 600
GAGTGGACAT GAGGCAAATT GAGAAGACAA TTCAATATTT GATTGGTTCA GGAATGGATC 660
CACGGACAGA AAACAGTCCA TACCTTGGGT TCATCTATAC ATCATTCAG GAAAGGGCAA 720
CCTTCATTTT TCATGGGAAC ACTGCCCCGAC AAGCCAAAGA GCATGGAGAC ATAAAGTTGG 780
CTCAAATATG TGGTACAATT GCTGCAGATG AGAAGCGCCA TGAGACAGCC TACACAAAGA 840
TAGTGGAAAA ACTCTTTGAG ATTGATCCTG ATGGAAGTGT TTTGGCTTTT GCTGATATGA 900
TGAGAAAGAA AATTTCTATG CCTGCACACT TGATGTATGA TGGCCGAGAT GATAATCTTT 960
TTGACCACTT TTCAGCTGTT GCGCAGCGTC TTGGAGTCTA CACAGCAAAG GATTATGCAG 1020
ATATATTGGA GTTCTTGGTG GGCAGATGGA AGGTGGATAA ACTAACGGGC CTTTCAGCTG 1080
AGGGACAAAA GGCTCAGGAC TATGTTTGTC GGTTACCTCC AAGAATTAGA AGGCTGGAAG 1140
```

AGAGAGCTCA AGGAAGGGCA AAGGAAGCAC CCACCATGCC TTTCAGCTGG ATTTTCGATA 1200
 GGCAAGTGAA GCTGTAGGTG GCTAAAGTGC AGGACGAAAC CGAAATGGTT AGTTTCACTC 1260
 TTTTTCATGC CCATCCCTGC AGAATCAGAA GTAGAGGTAG AATTTTGTAG TTGCTTTTTT 1320
 ATTACAAGTC CAGTTTAGTT TAAGGTCTGT GGAAGGGAGT TAGTTGAGGA GTGAATTTAG 1380
 TAAGTTGTAG ATACAGTTGT TTCTTGTTGT GTCATGAGTA TGCTGATAGA GAGCAGCTGT 1440
 AGTTTTGTTG TTGTGTTCTT TTATATGGTC TCTTGATGA GTTTCTTTTC TTTCTTTTC 1500
 TTCTTTCCTT TCCTCTCTCT CTCTCTCTCT CTCTCTCTTT TTCTCTTATC CCAAGTGTCT 1560
 CAAGTATAAT AAGCAAACGA TCCATGTGGC AATTTTGATG ATGGTGATCA GTCTCACAAC 1620
 TTGATCTTTT GTCTTCTATT GGAAACACAG CCTGCTTGTT TGAAAAAA 1668

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 396 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

MET	Ala	Leu	Lys	Leu	Asn	Pro	Phe	Leu	Ser	Gln	Thr	Gln	Lys	Leu	Pro
1				5					10					15	
Ser	Phe	Ala	Leu	Pro	Pro	MET	Ala	Ser	Thr	Arg	Ser	Pro	Lys	Phe	Tyr
			20					25					30		
MET	Ala	Ser	Thr	Leu	Lys	Ser	Gly	Ser	Lys	Glu	Val	Glu	Asn	Leu	Lys
		35					40					45			
Lys	Pro	Phe	MET	Pro	Pro	Arg	Glu	Val	His	Val	Gln	Val	Thr	His	Ser
	50					55					60				
MET	Pro	Pro	Gln	Lys	Ile	Glu	Ile	Phe	Lys	Ser	Leu	Asp	Asn	Trp	Ala
65					70				75						80
Glu	Glu	Asn	Ile	Leu	Val	His	Leu	Lys	Pro	Val	Glu	Lys	Cys	Trp	Gln
			85					90						95	

Pro	Gln	Asp	Phe	Leu	Pro	Asp	Pro	Ala	Ser	Asp	Gly	Phe	Asp	Glu	Gln	
			100					105					110			
Val	Arg	Glu	Leu	Arg	Glu	Arg	Ala	Lys	Glu	Ile	Pro	Asp	Asp	Tyr	Phe	
		115					120					125				
Val	Val	Leu	Val	Gly	Asp	MET	Ile	Thr	Glu	Glu	Ala	Leu	Pro	Thr	Tyr	
	130					135					140					
Gln	Thr	MET	Leu	Asn	Thr	Leu	Asp	Gly	Val	Arg	Asp	Glu	Thr	Gly	Ala	
145					150					155					160	
Ser	Pro	Thr	Ser	Trp	Ala	Ile	Trp	Thr	Arg	Ala	Trp	Thr	Ala	Glu	Glu	
				165					170					175		
Asn	Arg	His	Gly	Asp	Leu	Leu	Asn	Lys	Tyr	Leu	Tyr	Leu	Ser	Gly	Arg	
			180					185					190			
Val	Asp	MET	Arg	Gln	Ile	Glu	Lys	Thr	Ile	Gln	Tyr	Leu	Ile	Gly	Ser	
		195					200					205				
Gly	MET	Asp	Pro	Arg	Thr	Glu	Asn	Ser	Pro	Tyr	Leu	Gly	Phe	Ile	Tyr	
	210					215					220					
Thr	Ser	Phe	Gln	Glu	Arg	Ala	Thr	Phe	Ile	Ser	His	Gly	Asn	Thr	Ala	
225					230					235					240	
Arg	Gln	Ala	Lys	Glu	His	Gly	Asp	Ile	Lys	Leu	Ala	Gln	Ile	Cys	Gly	
				245					250					255		
Thr	Ile	Ala	Ala	Asp	Glu	Lys	Arg	His	Glu	Thr	Ala	Tyr	Thr	Lys	Ile	
			260					265					270			
Val	Glu	Lys	Leu	Phe	Glu	Ile	Asp	Pro	Asp	Gly	Thr	Val	Leu	Ala	Phe	
		275					280					285				
Ala	Asp	MET	MET	Arg	Lys	Lys	Ile	Ser	MET	Pro	Ala	His	Leu	MET	Tyr	
	290					295					300					
Asp	Gly	Arg	Asp	Asp	Asn	Leu	Phe	Asp	His	Phe	Ser	Ala	Val	Ala	Gln	
305					310					315					320	
Arg	Leu	Gly	Val	Tyr	Thr	Ala	Lys	Asp	Tyr	Ala	Asp	Ile	Leu	Glu	Phe	
				325					330					335		
Leu	Val	Gly	Arg	Trp	Lys	Val	Asp	Lys	Leu	Thr	Gly	Leu	Ser	Ala	Glu	
			340					345					350			
Gly	Gln	Lys	Ala	Gln	Asp	Tyr	Val	Cys	Arg	Leu	Pro	Pro	Arg	Ile	Arg	
		355					360					365				
Arg	Leu	Glu	Glu	Arg	Ala	Gln	Gly	Arg	Ala	Lys	Glu	Ala	Pro	Thr	MET	
	370					375					380					

Pro Phe Ser Trp Ile Phe Asp Arg Gln Val Lys Leu
 385 390 395

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 117 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TGAGAGATAG TGTGAGAGCA TTAGCCTTAG AGAGAGAGAG AGAGAGCTTG TGTCTGAAAG 60
 AATCCACAA ATG GCA TTG AAG CTT AAC CCT TTG GCA TCT CAG CCT TAC AAC 111
 MET Ala Leu Lys Leu Asn Pro Leu Ala Ser Gln Pro Tyr Asn
 1 5 10
 TTC CCT 117
 Phe Pro
 15

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 176 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ACT TCA TGG GCT ATT TGG ACA AGA GCT TGG ACT GCA GAA GAG AAC CGA 48
 Thr Ser Trp Ala Ile Trp Thr Arg Ala Trp Thr Ala Glu Glu Asn Arg
 1 5 10 15

CAC	GGT	GAT	CTT	CTC	AAT	AAG	TAT	CTT	TAC	TTG	TCT	GGA	CGT	GTT	GAC	96
His	Gly	Asp	Leu	Leu	Asn	Lys	Tyr	Leu	Tyr	Leu	Ser	Gly	Arg	Val	Asp	
			20					25					30			
ATG	AGG	CAG	ATT	GAA	AAG	ACC	ATT	CAG	TAC	TTG	ATT	GGT	TCT	GGA	ATG	144
MET	Arg	Gln	Ile	Glu	Lys	Thr	Ile	Gln	Tyr	Leu	Ile	Gly	Ser	Gly	MET	
		35					40					45				
GAT	CCT	AGA	ACA	GAG	AAC	AAT	CCT	TAC	CTC	GG						176
Asp	Pro	Arg	Thr	Glu	Asn	Asn	Pro	Tyr	Leu							
	50					55										

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1495 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGAGAGATAG	TGTGAGAGCA	TTAGCCTTAG	AGAGAGAGAG	AGAGAGCTTG	TGTCTGAAAG	60
AATCCACAAA	TGGCATTGAA	GCTTAACCCT	TTGGCATCTC	AGCCTTACAA	CTTCCCTTCC	120
TCGGCTCGTC	CGCCAATCTC	TACTTTCAGA	TCTCCCAAGT	TCCTCTGCCT	CGCTTCTTCT	180
TCTCCCGCTC	TCAGCTCCAA	GGAGGTTGAG	AGTTTGAAGA	AGCCATTAC	ACCACCTAAG	240
GAAGTGCACG	TTCAAGTCCT	GCATTCCATG	CCACCCCAGA	AGATCGAGAT	CTTCAAATCC	300
ATGGAAGACT	GGGCCGAGCA	GAACCTTCTA	ACTCAGCTCA	AAGACGTGGA	GAAGTCGTGG	360
CAGCCCCAGG	ACTTCTTACC	CGACCCTGCA	TCCGATGGGT	TCGAAGATCA	GGTTAGAGAG	420
CTAAGAGAGA	GGGCAAGAGA	GCTCCCTGAT	GATTACTTCG	TTGTTCTGGT	GGGAGACATG	480
ATCACGGAAG	AGGCGCTTCC	GACCTATCAA	ACCATGTTGA	ACACTTTGGA	TGGAGTGAGG	540
GATGAAACTG	GCGCTAGCCC	CACTTCATGG	GCTATTTGGA	CAAGAGCTTG	GA CTGCAGAA	600
GAGAACCGAC	ACGGTGATCT	TCTCAATAAG	TATCTTTACT	TGTCTGGACG	TGTTGACATG	660
AGGCAGATTG	AAAAGACCAT	TCAGTACTTG	ATTGGTTCTG	GAATGGATCC	TAGAACAGAG	720

AACAAATCCTT ACCTCGGCTT CATCTACACT TCATTCCAAG AAAGAGCCAC CTTCATCTCT 780
 CACGGAAACA CAGCTCGCCA AGCCAAAGAG CACGGAGACC TCAAGCTAGC CCAAATCTGC 840
 GGCACAATAG CTGCAGACGA GAAGCGTCAT GAGACAGCTT ACACCAAGAT AGTTGAGAAG 900
 CTCTTTGAGA TTGATCCTGA TGGTACTGTG ATGGCGTTTG CAGACATGAT GAGGAAGAAA 960
 ATCTCGATGC CTGCTCACTT GATGTACGAT GGGCGGGATG AAAGCCTCTT TGACAACCTC 1020
 TCTTCTGTTG CTCAGAGGCT CGGTGTTTAC ACTGCCAAAG ACTATGCGGA CATTCTTGAG 1080
 TTTTGTGTTG GGAGGTGGAA GATTGAGAGC TTGACCGGGC TTTCAGGTGA AGGAAACAAA 1140
 GCGCAAGAGT ACTTGTGTGG GTTGA CTCCA AGAATCAGGA GGTGATGA GAGAGCTCAA 1200
 GCAAGAGCCA AGAAAGGACC CAAGGTTCTT TTCAGCTGGA TACATGACAG AGAAGTGCAG 1260
 CTCTAAAAAG GAACAAAGCT ATGAAACCTT TTCACTCTCC GTCGTCCCTC ATTTGATCTA 1320
 TCTGCTCTTG AAATTGGTGT AGATTACTAT GGTGTTGTGAT ATTGTTTCGTG GGTCTAGTTA 1380
 CAAAGTTGAG AAGCAGTGAT TTAGTAGCTT TGTTGTTTCC AGTCTTTAAA TGTTTTTGTG 1440
 TTTGGTCCTT TTAGTAAACT TGTTGTAGTT AAATCAGTTG AACTGTTTGG TCTGT 1495

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 398 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

MET	Ala	Leu	Lys	Leu	Asn	Pro	Leu	Ala	Ser	Gln	Pro	Tyr	Asn	Phe	Pro
1				5					10					15	
Ser	Ser	Ala	Arg	Pro	Pro	Ile	Ser	Thr	Phe	Arg	Ser	Pro	Lys	Phe	Leu
			20					25					30		
Cys	Leu	Ala	Ser	Ser	Ser	Pro	Ala	Leu	Ser	Ser	Lys	Glu	Val	Glu	Ser
		35					40					45			
Leu	Lys	Lys	Pro	Phe	Thr	Pro	Pro	Lys	Glu	Val	His	Val	Gln	Val	Leu
50						55					60				

His 65	Ser	MET	Pro	Pro	Gln 70	Lys	Ile	Glu	Ile	Phe 75	Lys	Ser	MET	Glu	Asp 80
Trp	Ala	Glu	Gln	Asn 85	Leu	Leu	Thr	Gln	Leu 90	Lys	Asp	Val	Glu	Lys 95	Ser
Trp	Gln	Pro	Gln 100	Asp	Phe	Leu	Pro	Asp 105	Pro	Ala	Ser	Asp	Gly 110	Phe	Glu
Asp	Gln	Val 115	Arg	Glu	Leu	Arg	Glu 120	Arg	Ala	Arg	Glu	Leu 125	Pro	Asp	Asp
Tyr 130	Phe	Val	Val	Leu	Val	Gly 135	Asp	MET	Ile	Thr	Glu 140	Glu	Ala	Leu	Pro
Thr 145	Tyr	Gln	Thr	MET	Leu 150	Asn	Thr	Leu	Asp	Gly 155	Val	Arg	Asp	Glu	Thr 160
Gly	Ala	Ser	Pro	Thr 165	Ser	Trp	Ala	Ile	Trp 170	Thr	Arg	Ala	Trp	Thr 175	Ala
Glu	Glu	Asn 180	Arg	His	Gly	Asp	Leu 185	Leu	Asn	Lys	Tyr	Leu 190	Tyr	Leu	Ser
Gly	Arg	Val 195	Asp	MET	Arg	Gln	Ile 200	Glu	Lys	Thr	Ile	Gln 205	Tyr	Leu	Ile
Gly	Ser 210	Gly	MET	Asp	Pro	Arg 215	Thr	Glu	Asn	Asn	Pro 220	Tyr	Leu	Gly	Phe
Ile 225	Tyr	Thr	Ser	Phe	Gln 230	Glu	Arg	Ala	Thr	Phe 235	Ile	Ser	His	Gly	Asn 240
Thr	Ala	Arg	Gln	Ala 245	Lys	Glu	His	Gly	Asp 250	Leu	Lys	Leu	Ala	Gln 255	Ile
Cys	Gly	Thr 260	Ile	Ala	Ala	Asp	Glu 265	Lys	Arg	His	Glu	Thr 270	Ala	Tyr	Thr
Lys	Ile 275	Val	Glu	Lys	Leu	Phe	Glu 280	Ile	Asp	Pro	Asp	Gly 285	Thr	Val	Met
Ala 290	Phe	Ala	Asp	MET	MET	Arg 295	Lys	Lys	Ile	Ser	Met 300	Pro	Ala	His	Leu
Met 305	Tyr	Asp	Gly	Arg	Asp 310	Glu	Ser	Leu	Phe	Asp 315	Asn	Phe	Ser	Ser	Val 320
Ala	Gln	Arg	Leu	Gly 325	Val	Tyr	Thr	Ala	Lys 330	Asp	Tyr	Ala	Asp	Ile 335	Leu
Glu	Phe	Leu	Val 340	Gly	Arg	Trp	Lys	Ile 345	Glu	Ser	Leu	Thr	Gly 350	Leu	Ser

Gly	Glu	Gly	Asn	Lys	Ala	Gln	Glu	Tyr	Leu	Cys	Gly	Leu	Thr	Pro	Arg
		355					360					365			
Ile	Arg	Arg	Leu	Asp	Glu	Arg	Ala	Gln	Ala	Arg	Ala	Lys	Lys	Gly	Pro
	370					375					380				
Lys	Val	Pro	Phe	Ser	Trp	Ile	His	Asp	Arg	Glu	Val	Gln	Leu		
385					390					395					

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCTAAGCTTA ARGARATHCC AGAYGAYTA 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTAAGCTTA ARGARATHCC GGAYGAYTA 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCTAAGCTTA ARGARATHCC CGAYGAYTA 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCTAAGCTTA ARGARATHCC TGAYGAYTA 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

AGCGAATTCG TRTTNAGCAT NGTYTG 26

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide mixture

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AGCGAATTCG TRTTYAACAT NGTYTG 26

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3440 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TCTAGAATTC TCTAATTACG TCTGTTTGTT CTATTTTTTA TATGATATCA AATATTCGTC 60
ATAAATATAT GGTTTAAGAT GCCAAAAAAT TATTTACTTG GTGAATATAA TACGTAAAT 120
ATTAGAAATA CATCATTTAG TTAAATAAAT AACCAAAAAC CAAAATTCA TATCCGCGCT 180
GGCGCGCGGT CAGGGTCTCG TTAGTTTTAA AATCAATGCA GTTTACAATT AATTTCAGC 240

TGAAAATAAG	TATAATTTGT	ATTGAAATTA	TAAAGTGACA	TTTTTTGTGT	AACAAATATT	300
TTGTGTAACA	AGAATTAAAA	AAAAAACAG	AAAATACTCA	GCTTTTTTAA	TAATAAAAAA	360
AATTAATTGA	GTTAGAAAAT	TGTTGTACCA	ATAACAAAAG	ATTTATATGG	AATTATAAAA	420
TCAACACACC	AATAACACAA	GACTTTTTTAA	AAATTTAAGA	ATAATATAAG	CAATAACAAT	480
AGAATCTTCA	AATTCTTCAA	ATCCTTAAAA	ATCAATCTCC	CACTATTAAT	CCCCCTTAGT	540
TTTAGTTGGT	AATGGCAACG	TTTGTTGACT	ACCGTATTGT	AAC TTTTGTC	AAATTGTCAT	600
AAATACGTGT	CAAACCTCTGG	TAAAAAATTA	GTCTGCTACA	TCTGTCTTTT	ATTTATAAAA	660
CACAGCTGTT	AATCAGAATT	TGGTTTATTA	AATCAACAAC	CTGCACGAAA	CTTGTGTGAG	720
CATATTTTGT	CTGTTTCTGG	TTCATGACCT	TCTTCCGCAT	GATGGCCAAG	TGTAATGGCC	780
ACTTGCAAGA	GCGTTTCTTC	AACGAGATAA	GTCGAACAAA	TATTTGTCCG	TTACGACCAC	840
ATATAAAATC	TCCCCATCTC	TATATATAAT	ACCAGCATTG	ACCATCATGA	ATACCTCAAA	900
TCCAATCTC	ACAAATACTT	CAATAAAAAG	ACCAAAAAAA	ATTAAAGCAA	AGAAAAGCCT	960
TCTTGTGCAC	AAAAAAAAAA	GAAGCCTTCT	AGGTTTTTAC	GACATGAAGT	TCACTACTCT	1020
AATGGTCATC	ACATTGGTGA	TAATCGCCAT	CTCGTCTCCT	GTTCCAATTA	GAGCAACCAC	1080
GGTTGAAAGT	TTCGGAGAAG	TGGCACAATC	GTGTGTTGTG	ACAGAACTCG	CCCCATGCTT	1140
ACCAGCAATG	ACCACGGCAG	GAGACCCGAC	TACAGAATGC	TGCGACAAAC	TGGTAGAGCA	1200
GAAACCATGT	CTTTGTGGTT	ATATTCGAAA	CCCAGCCTAT	AGTATGTATG	TTACTTCTCC	1260
AAACGGTCGC	AAAGTCTTAG	ATTTTTGTAA	GGTTCCTTTT	CCTAGTTGTT	AAATCTCTCA	1320
AGACATTGCT	AAGAAAAATA	TTATTAAAAA	TAAAAGAATC	AAACTAGATC	TGATGTAACA	1380
ATGAATCATC	ATGTTATGGT	TGAAGCTTAT	ATGCTGAAGT	GTTTGATTTT	ATATATGTGT	1440
GTGTGTGTGT	CCTGCTCAAT	TTTTGAAACA	CACACGTTTC	TCCTGATTTG	GATTTAAATT	1500
ATATTTTGAG	TTAAAAAAA	GAAAAAGATG	GAATGCTATT	TATACAAGTT	GATGAAAAAG	1560
TGGAAGTACA	ATTTAGATAT	CTCCTACACT	TAAAGAATGA	AACAATAATA	GACTTACGAA	1620
ACAAATGAAA	AATACATAAA	TTGTGACAAA	TCAACGTCCG	ATGACGAGTT	TATTATTAAA	1680
AATTTGTGTG	AAGGACTAGC	AGTTCAACCA	AATGATATTG	AACATATACA	TCAACAAATA	1740
TGATAATCAT	AAAAGAGAGA	ATGGGGGGGG	GGTGTGTTTT	ACCAGAAACC	TCTTTTTTCTC	1800
AGCTCGCTAA	AACCCTACCA	CTAGAGACCT	AGCTCTGACC	GTCGGCTCAT	CGGTGCCGGA	1860

GGTGTAACCT	TTCTTTCCCA	TGACCCGAAA	CCTCTCTTTC	CCAACTCACG	AAAACCCTAC	1920
AATCAAAAAC	CTAGCTCCGA	CCGTGGGCTC	ATCGGTGCCG	AAGGTGTAAC	CTTTCTCTCC	1980
CATCATAGTT	TCTCGTAAAT	GAAAGCTAAT	TGGGCAATCG	ATTTTTTAAT	GTTTAAACCA	2040
TGCCAAGCCA	TTTCTTATAG	GACAATTGTC	AATAATAGCA	TCTTTTGAGT	TTTGTCTCAA	2100
AAGTGACACT	AGAAGAAAAA	AGTCACAAAA	ATGACATTCA	TTAAAAAGTA	AAATATCCCT	2160
AATACCTTTG	GTTTAAATTA	AATAAGTAAA	CAAAAATAAA	TAAAAACAAA	TAAAATAAAA	2220
ATAAAAAATG	AAAAAAAGAA	ATTTTTTTTAT	AGTTTCAGAT	TATATGTTTT	CAGATTCGAA	2280
ATTTTTTAAA	TTCCCTTTTT	TAAATTTTCT	TTTTTGAAAT	TTTTTTTTTT	GAAATTTTTT	2340
GAAACTGTTT	TTAAAATTTT	TATTTTTAAT	TTTTTAGTAT	TTATTTTTTA	TTTTATAAAA	2400
TTTTTAAACGC	TAATTCCAAA	ACTCCCCCCC	CCCCCCCCC	CCCAATTCTC	TCCTAGTCTT	2460
TTTCTCTTTC	TTATATTGG	GCTTCTATCT	TCTCTTTTTT	TTTCAGGCC	AAAGTATCAT	2520
GTGTAACAAC	CGGTGTTCAA	AAACGCGCCC	GCCTGGCCGT	TTACTCGCCC	GATTAAATGA	2580
TGATCGGAAG	GCTGCCATGG	CGAGGCGGAG	GTAATCAGTG	GTTCTAGGCG	CTGAAACTAG	2640
AAAACCTTCA	AAAATCGAAA	TTTTAAGAGC	TAAATCGGTG	TTTATCTCAT	GAATCTATTA	2700
TATTTAGTTG	AAACTCACAA	GAATCGGTTG	TAAAACTAT	GAAATCGTGC	AAAAAAAATG	2760
AAGAACAAAA	TATTCTCAGA	TCTGGAAAAC	ACAGAGAAGA	GGTTGAAGAT	GAGGGTAAAA	2820
TCGTATTTTG	TCATTCATTA	AACTAAAATC	AAAAAAAAT	GATGCAAAAT	TCAATGATAA	2880
TAACTCGAAC	TCGCAACCAT	ATGCATCTTT	AGACTGCGAC	ACGGACCACT	AGACTAAGCA	2940
ATTTTAATGT	TTATTCATCA	CAGACCTAAT	ATATGTCTAA	AACTAGGCGC	CGAGTACGCC	3000
CCGCTTAATC	CCGAGTTTTT	GTTAGCTCGC	TAGACCCAGG	GTCACCGCCC	GACTAACGAG	3060
TAGCGTAATT	CTGAACTGGG	GTAACAACAT	AGAGAACATC	GCCGACCCTT	CCCTGCCGAT	3120
GATGCCGCCT	CCGATGAACT	TCCTGTAACG	CCTTCAGTTT	CCATTGATTT	TCCCCTTTAA	3180
TCTGATCAGT	TCCATGTTTT	ATCCAACCTCA	TCCCCTCCG	TAGCATTTAA	TCGATCTCAT	3240
CATTTACATA	CATAACCAGT	AGGAGGTCTC	ATATAAATTT	GAACGTTTCC	AGCGATGAAC	3300
AGTGCCAATC	TCTGCGAAAT	CCATTTCTCT	AAGCTCAGGG	CTGGCGGCTG	CAGCCCGGGG	3360
ATCCACTAGT	TCTAGGCGGC	CGCACC GCGG	TGGAGCTCCA	ATTCGCCCTA	TAGTGAGTCG	3420
TATTACGCGC	GCTCACTGGC					3440

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3898 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CTCGAGAGCT	GAAGGATTTT	TTGTTAGAGA	TTCAACGACA	GATGGACCCT	TCCTCCACTA	60
GGCAACTGCA	AGAACCTAAC	AATGCAAATA	TCACTCCTCC	TCAGCCTTCA	AGGAGCGTTA	120
ATAGGACTGG	AACAAGCGGT	CAAGTGAGTA	AATTTTCCTT	CCAAGATAGA	TCTCTATGGT	180
TCGGTTCATG	AAGTTTGTGG	TTTAATTGTG	TAGCAACAGG	ATAGTGCAAG	TGAGAATAGA	240
GTTCGACCTC	ATCTACCTAC	CCCGGAACCT	CTGAATGTAT	CCCCATTGAA	GAAGAAGAGG	300
GCAAATCCTG	CACCCAGAAG	GATAAAGAAA	TTTTGGACGC	CTGAAGAAGT	GGCAGTTCTG	360
AGGGAAGGAG	TAAAAGAGTA	TGTCTACTAC	TACTACTCTA	TAATCAAGTT	TCAAGAAGCT	420
GAGCTTGGCT	CTCACTTTAT	ATGTTTGATG	TTGTTGTGCA	GGTATGGTAA	ATCATGGAAA	480
GAGATAAAGA	ATGCAAACCC	TGAAGTATTG	GCAGAGAGGA	CTGAGGTGAG	AGAGCATGTC	540
ACTTTTGTGT	TACTCATCTG	AATTATCTTA	TATGCGAATT	GTAAGTGGTA	CTAAAAGGTT	600
TGTAACCTTT	GGTAGGTGGA	TTTGAAGGAT	AAATGGAGGA	ACTTGCTTCG	GTAGCGGTAA	660
CAAGTTTTAT	ATTGCTATGA	AGTTTTTTTG	CCTGCGTGAC	GTATCAGCAG	CTGTGGAGAA	720
GATGGTATTA	GAAAGGGTCT	TTTCACATTT	TGTGTTGTGA	CAAATATTAA	TTCGGCCGGT	780
ATGGTTTGGT	TAAGACTTGT	TGAGAGACGT	GTGGGGTTTT	TTGATGTATA	ATTAGTCTGT	840
GTTTAGAACG	AAACAAGACT	TGTTGCGTAT	GCTTTTTTTA	ACTTGAGGGG	GTTTGTGTGT	900
GTTAGTTAGG	AACTTGACTT	TGTCTCTTTC	TCTCAAGATC	TGATTGGTAA	GGTCTGGGTG	960
GTAGTACTGT	TTGGTTTAAT	TTGTTTTGAC	TATTGAGTCA	CTGTGGCCCA	TTGACTTTAA	1020
ATTAGGCTGG	TATATTTTTT	GGTTTAAAC	CGGTCTGAGA	TAGTGCAATT	TCGATTCAGT	1080

CAATTTTAAA	TTCTTCAAGG	TAATGGGCTG	AATACTTGTA	TAGTTTTAAG	ACTTAACAGG	1140
CCTTAAAAGG	CCCATGTTAT	CATAAAACGT	CATTGTTTAG	AGTGCACCAA	GCTTATAAAA	1200
TGTAGCCAGG	CCTTAAAAGA	CTTAACAGGC	CTTAAAAGAC	TTAACATTCC	TTAAAAGGCC	1260
CATGTTATCA	TAAAACGTCA	TCGTTTTGAG	TGCACCAAGC	TAAATGTAGC	CAGGCCTTAA	1320
AAGACTTAAC	AGGCCTTAAA	AGGCCCATGT	TATCATAAAA	CGCCGTCGTT	TTGAGTGCAC	1380
CAAGCTTATA	AATGTAGCCA	GCTACCTCGG	GACATCACGC	TCTTTGTACA	CTCCGCCATC	1440
TCTCTCTCTC	TCGAGCAGAT	CTCTCTCGGG	AATATCGACA	ATGTCGACCA	CTTTCTGCTC	1500
TTCCGTCCTC	ATGCAAGCCA	CTTCTCTGGT	AATCTCATCT	CCTTCTTGTTG	TTCCCAGATC	1560
GCTCTGATCA	TACTTTCTTT	TAGATCATTT	GCCTCTGATC	TGTTGCTTGA	TGTTTGTTAA	1620
CTCTCCACGC	ATGTTTGATT	ATGTTGAGAA	TTAGAAAAAA	AATGTTAGCT	TTACGAATCT	1680
TTAGTGATCA	TTTCAATTGG	ATTTGCAATC	TTGTGTGACA	TTTGAGGCTT	GTGTAGATTT	1740
CGATCTGTAT	TCATTTTGAA	TCACAGCTAT	AATAGTCATT	TGAGTAGTAG	TGTTTTTAAA	1800
TGAACATGTT	TTGTTGTATT	GATGGAACAA	ACAGGCAGCA	ACAACGAGGA	TTAGTTTCCA	1860
GAAGCCAGCT	TTGGTTTCAA	CGACTAATCT	CTCCTTCAAC	CTCCGCCGTT	CAATCCCCAC	1920
TCGTTTCTCA	ATCTCCTGCG	CGGTATGTTC	TCATTCTCAG	CATTTATTTT	GAGCTTGCTT	1980
GTCATGGTAC	TCTCTCTAAT	TGTCTATTTG	GTTTATTAGG	CCAAACCAGA	GACGGTTGAG	2040
AAAGTGTCTA	AGATAGTTAA	GAAGCAGCTA	TCACTCAAAG	ACGACCAAAA	GGTCGTTGCG	2100
GAGACCAAGT	TTGCTGATCT	TGGAGCAGAT	TCTCTCGACA	CTGTAAGTCA	TCAATCATTC	2160
TCTTATGTGA	ATAAAGAGAA	CTTGAAGAGT	TTGTTTTTAA	CATATTAAGT	GAGTGTTTTG	2220
CATGCAGGTT	GAGATAGTGA	TGGGTTTAGA	GGAAGAGTTT	GATATCGAAA	TGGCTGAAGA	2280
GAAAGCTCAG	AAGATTGCTA	CTGTGGAGGA	AGCTGCTGAA	CTCATTGAAG	AGCTCGTTCA	2340
ACTTAAGAAG	TAATTTTAGT	ATTAAGAGCA	GCCAAGGCTT	TGTTGGGTTT	GTTGTTTTCA	2400
TAATCTTCCT	GTCATTTTCT	TTTTCTTTAA	TGTGTCAAGC	GACTCTGTTG	GTTTAAAGTA	2460
GTATCTGTTT	GCCATGGATC	TCTCTCTATT	TGTCGACTGA	AACTTTTGG	TTTACACATG	2520
AAAGCTTGTT	CTTGTTCTTT	CTTAAATCGA	AATGCCAAAT	GCGAGATTAG	GGAATCTTGT	2580
ATTAACACAT	ACATAAGTCA	AAGAGTAGGC	CCTAAGATGA	CAATTTATAA	ACAATCCTAT	2640
TCACATTGTA	TATACAGGTT	ATGATTATTC	CCAATCAGCG	TCAAAGAATC	CAGCATCTTT	2700

CATCTCTGAA	TAGTAGACAT	TCTCCAAGTT	CACATCTTCC	TCCTGCACCA	AAAACCAGTA	2760
CTAAATCATG	AACATTGCAA	TAATCACATG	CCTAGGCGAG	AGTTTTGGTG	ATGTGGTGTT	2820
AGTGATAGTG	ATACTGATGG	TGCTAGAGCG	GTTAAGAAGG	ATTAACCTGG	AAGAAGTCTG	2880
CAAGGAAAGT	AACATAGAGA	AGAGGAAGAT	AGGAGTGGTA	ACAAACACTT	GTGATCCCAT	2940
ACAGCCTCCC	AGCATTTTTTC	AAATGTTATT	TCCTTACATA	AAGAAACAAG	AGAAGTCTGA	3000
CTAGATGATA	TTTATATAGG	ATAAGTGTTT	TACCATAAGC	CAAAGTGAGC	GCCGTTTGCA	3060
AGAGCTAACC	AGACAGTACA	CGTTTGGCAT	ATATCTCATC	AACATGATCT	GAAAAGTAAC	3120
ATATCACAGT	TAATGAACAC	AATGGTTACC	TTGAGAAGCA	AATCAAGACC	TATAACAAGC	3180
CCAGAGATGA	GGAAAGTCCG	TGTCAACGCT	TCACCGCCAT	TCGCGTAGTT	TCCTTGGAAG	3240
ACAAAGGCCA	CCAACCAAAC	TTACTTCCAG	AAACAACACT	CCAAATGTTG	TCAACAAAGT	3300
CAATAGATTC	CAAACACTT	CGTTACAGGG	TTGTATAGAT	AATATAATAG	AATAGTGGGA	3360
AGATAGTATA	AATAAAATAA	ATAAAAGATC	CTATCGGTAA	ATAGTTTATA	ATATCGGGGG	3420
CGTATATAAA	GTATAAAAGA	AACTCTTCTC	CAATCCGACC	GTTGAAAATC	ACTCTCAATC	3480
TCTGGCGTAA	CGACCGGATC	GTTCGCGCGT	AATTTTCGCT	GCTATAAATA	GAAACTTTCC	3540
TCTTCTGTTT	CTCGATCAAA	ATTTTTTTTT	GGAAAAATTA	AGTTTGAATC	TATCGTAGAT	3600
GCTGTGACAA	AAAAAAATTG	TTTTATCGAA	GATGAGAAAC	ATGAGGCCTG	TTCATGCAAG	3660
GAACCAGACC	ACGGATCCAT	CTTCGCCGAT	GATGACGTCT	CCTCTGATGA	ATCGTCACGC	3720
ACGGACAGGA	TCCAACGCTG	GACCAGCATC	TAACGCCAAG	AAAGCACAGA	CGAAAGCAGC	3780
AGCTCAGAGA	CTCGCGGCTG	TGATGTCGAA	CCAAACAGGC	GACGATGAAG	ACAGTGATGA	3840
TGACCTTTCC	TTTGA CTACA	ACGCTGTCGG	AAGCATTGGT	CTCGCTGCCG	GAAGATCT	3898

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4325 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CTCGAGGCAG	TCACTAACAT	GAAGTTTGAC	GAGGAGCCCA	ACTATGGGAA	GCTTATTTCT	60
CTTTTCGATA	CTCTAATTGA	GCCGTGCGCT	CTATCTAGAC	CAATTAGAAT	TGATGGAGCT	120
CTAAAGGTTG	CTGGCTGTTT	TCTTGTTTCAT	ATGATTAACT	TCTAAACTTG	TGTATAAATA	180
TTCTCTGAAA	GTGCTTCTTT	TGGCATATGT	AGGTTGGGCA	AAAACGAGGA	AGATTGCTTC	240
TCAATTTGGA	AGAGGATGAA	CAGCCGAAGA	AGAAAATAAG	AATAGGCAGT	CCTGCTACTC	300
AATGGATCTC	AGTCTATAAC	GGTCGTCGTC	CCATGAAACA	GAGGTAAAC	ATTTTTTGCA	360
TATACACTTT	GAAAGTTCCT	CACTAACTGT	GTAATCTTTT	GGTAGATATC	ACTACAATGT	420
CGGAGAGACA	ANGGCTGSNC	ANCATATACA	AAAGGGAAAT	GAAGATGGCC	TTTTGATTAG	480
CTGTGTAGCA	TCAGCAGCTA	ATCTCTGGGC	TCTCATCATG	GATGCTGGAA	CTGGATTAC	540
TTCTCAAGTT	TATGAGTTGT	CACCGGTCTT	CCTACACAAG	GTAATAATCA	GTTGAAGCAA	600
TTAAGAATCA	ATTTGATTTG	TAGTAAACTA	AGAAGAACTT	ACCTTATGTT	TTCCCCGCAG	660
GACTGGATTA	TGGAACAATG	GGAAAAGAAC	TACTATATAA	GCTCCATAGC	TGGTTCAGAT	720
AACGGGAGCT	CTTTAGTTGT	TATGTCAAAA	GGTTAGTGTT	TAGTGAATAA	TAAACTTATA	780
CCACAAAGTC	TTCATTGACT	TATTTATATA	CTTGTTGTGA	ATTGCTAGGA	ACTACTTATT	840
CTCAGCAGTC	ATACAAAGTG	AGTGACTCAT	TTCCGTTCAA	GTGGATAAAT	AAGAAATGGA	900
AAGAAGATTT	TCATGTAACC	TCCATGACAA	CTGCTGGTAA	TCGTTGGGGT	GTGGTAATGT	960
CGAGGAACTC	TGGCTTCTCT	GATCAGGTAG	GTTTTTGTCT	CTTATTGTCT	GGTGTTTTTA	1020
TTTTCCCCTG	ATAGTCTAAT	ATGATAAACT	CTGCGTTGTG	AAAGGTGGTG	GAGCTTGACT	1080
TTTTGTACCC	AAGCGATGGG	ATACATAGGA	GGTGGGAGAA	TGGGTATAGA	ATAACATCAA	1140
TGGCAGCAAC	TGCGGATCAA	GCAGCTTTCA	TATTAAGCAT	ACCAAAGCGT	AAGATGGTGG	1200
ATGAAACTCA	AGAGACTCTC	CGCACCACCG	CCTTTCCAAG	TACTCATGTC	AAGGTTGGTT	1260
TCTTTAGCTT	TGAACACAGA	TTTGGATCTT	TTTGTTTTGT	TTCCATATAC	TTAGGACCTG	1320
AGAGCTTTTG	GTTGATTTTT	TTTTCAGGAC	AAATGGGCGA	AGAATCTGTA	CATTGCATCA	1380
ATATGCTATG	GCAGGACAGT	GTGCTGATAC	ACACTTAAGC	ATCATGTGGA	AAGCCAAAGA	1440
CAATTGGAGC	GAGACTCAGG	GTCGTCATAA	TACCAATCAA	AGACGTAAAA	CCAGACGCAA	1500

CCTCTTTGGT TGAATGTAAT GAAAGGGATG TGTCTTGGTA TGTATGTACG AATAACAAAA 1560
GAGAAGATGG AATTAGTAGT AGAAATATTT GGGAGCTTTT TAAGCCCTTC AAGTGTGCTT 1620
TTTATCTTAT TGATATCATC CATTTGCGTT GTTTAATGCG TCTCTAGATA TGTTCCTATA 1680
TCTTTCTCAG TGTCTGATAA GTGAAATGTG AGAAAACCAT ACCAAACCAA AATATTCAAA 1740
TCTTATTTTT AATAATGTTG AATCACTCGG AGTTGCCACC TTCTGTGCCA ATTGTGCTGA 1800
ATCTATCACA CTAGAAAAAA ACATTTCTTC AAGGTAATGA CTTGTGGACT ATGTTCTGAA 1860
TTCTCATTA GTTTTTATTT TCTGAAGTTT AAGTTTTTAC CTTCTGTTTT GAAATATATC 1920
GTTCATAAGA TGTCACGCCA GGACATGAGC TACACATCGC ACATAGCATG CAGATCAGGA 1980
CGATTTGTCA CTCACTTCAA ACACCTAAGA GCTTCTCTCT CACAGCGCAC ACACATATGC 2040
ATGCAATATT TACACGTGAT CGCCATGCAA ATCTCCATTC TCACCTATAA ATTAGAGCCT 2100
CGGCTTCACT CTTTACTCAA ACCAAAACCTC ATCACTACAG AACATACACA AATGGCGAAC 2160
AAGCTCTTCC TCGTCTCGGC AACTCTCGCC TTGTTCTTCC TTCTCACCAA TGCCTCCGTC 2220
TACAGGACGG TTGTGGAAGT CGACGAAGAT GATGCCACAA ATCCAGCCGG CCCATTTAGG 2280
ATTCCAAAAT GTAGGAAGGA GTTTCAGCAA GCACAACACC TGAAAGCTTG CCAACAATGG 2340
CTCCACAAGC AGGCAATGCA GTCCGGTAGT GGTCCAAGCT GGACCCTCGA TGGTGAGTTT 2400
GATTTTGAAG ACGACGTGGA GAACCAACAA CAGGGCCCCG AGCAGAGGCC ACCGCTGCTC 2460
CAGCAGTGCT GCAACGAGCT CCACCAGGAA GAGCCACTTT GCGTTTGCCC AACCTTGAAA 2520
GGAGCATCCA AAGCCGTAA ACAACAGATT CGACAACAAC AGGGACAACA AATGCAGGGA 2580
CAGCAGATGC AGCAAGTGAT TAGCCGTATC TACCAGACCG CTACGCACTT ACCTAGAGCT 2640
TGCAACATCA GGCAAGTTAG CATTTGCCCC TTCCAGAAGA CCATGCCTGG GCCCGGCTTC 2700
TACTAGATTC CAAACGAATA TCCTCGAGAG TGTGTATACC ACGGTGATAT GAGTGTGGTT 2760
GTTGATGTAT GTTAACACTA CATAGTCATG GTGTGTGTTC CATAAATAAT GTACTAATGT 2820
AATAAGAACT ACTCCGTAGA CGGTAATAAA AGAGAAGTTT TTTTTTTTAC TCTTGCTACT 2880
TTCCTATAAA GTGATGATTA ACAACAGATA CACCAAAAAG AAAACAATTA ATCTATATTC 2940
ACAATGAAGC AGTACTAGTC TATTGAACAT GTCAGATTTT CTTTTTCTAA ATGTCTAATT 3000
AAGCCTTCAA GGCTAGTGAT GATAAAAGAT CATCCAATGG GATCCAACAA AGACTCAAAT 3060
CTGGTTTTGA TCAGATACTT CAAAACATTT TTTGTATTCA TTAAATTATG CAAGTGTCTT 3120

TTTATTTGGT GAAGACTCTT TAGAAGCAAA GAACGACAAG CAGTAATAAA AAAAACAAG 3180
 TTCAGTTTTA AGATTTGTTA TTGACTTATT GTCATTTGAA AAATATAGTA TGATATTAAT 3240
 ATAGTTTAT TTATATAATG CTTGTCTATT CAAGATTTGA GAACATTAAT ATGATACTGT 3300
 CCACATATCC AATATATTAA GTTTCATTTT TGTTCAAACA TATGATAAGA TGGTCAAATG 3360
 ATTATGAGTT TTGTTATTTA CCTGAAGAAA AGATAAGTGA GCTTCGAGTT TCTGAAGGGT 3420
 ACGTGATCTT CATTTCTTGG CTAAAAGCGA ATATGACATC ACCTAGAGAA AGCCGATAAT 3480
 AGTAAACTCT GTTCTTGGTT TTTGGTTTAA TCAAACCGAA CCGGTAGCTG AGTGTCAAGT 3540
 CAGCAAACAT CGCAAACCAT ATGTCAATTC GTTAGATTCC CGGTTTAAAGT TGTAACCCGG 3600
 TATTTTATTT GGTGAAAACC CTAGAAGCCA GCCANCCTTT TTAATCTAAT TTTTGCAAAC 3660
 GAGAAGTCAC CACACCTCTC CACTAAAACC CTGAACCTTA CTGAGAGAAG CAGAGNCANN 3720
 AAAGAACAAA TAAAACCCGA AGATGAGACC ACCACGTGCG GCGGGACGTT CAGGGGACGG 3780
 GGAGGAAGAG AATGRCGGCG GNSNTTTGGT GGCGGCGGCG GACGTTTTTG TGGCGGCGGT 3840
 GGACGTTTTG GTGGCGGCGG TGGACCTTTG GTGGTGGATA TCGTGACGAA GGACCTCCCA 3900
 GTGAAGTCAT TGGTTCGTTT ACTCTTTTCT TAGTCGAATC TTATTCTTGC TCTGCTCGTT 3960
 GTTTTACCGA TAAAGCTTAA GACTTTATTG ATAAAGTTCT CAGCTTTGAA TGTGAATGAA 4020
 CTGTTTCCTG CTTATTAGTG TTCCTTTGTT TTGAGTTGAA TCACTGTCTT AGCACTTTTG 4080
 TTAGATTCAT CTTTGTGTTT AAGTTAAAAG GTAGAACTT TGTGACTTGT CTCCGTTATG 4140
 ACAAGGTAA CTTTGTGTTT TATAACAGAA GTTGCGACCT TTCTCCATGC TTGTGAGGGT 4200
 GATGCTGTGG ACCAAGCTCT CTCAGGCGAA GATCCCTTAC TTCAATGCCC CAATCTACTT 4260
 GGAAAACAAG ACACAGATTG GGAAAGTTGA TGAGATCCAA GCTTGGGCTG CAGGTCGACG 4320
 AATTC 4325

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGGATCCACT GCAGTCTAGA GGGCCCGGGA 30

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AATTTCCCGG GCCCTCTAGA CTGCAGTGGA TCCGAGCT 38

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

GTAAGTAGGT AGGGCTTCCT CTGTAATCAT ATCTCCAACC AAAACAACAA 50

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTTAAGAAGT AACCCGGGCT GCAGTTT TAG TATTAAGAG 39

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GGAATTCGTC GACAGATCTC TGCAGCTCGA GGGATCCAAG CTT 43

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCATTTTGA TCTTCCTCGA GCCCGGGCTG CAGTTCTTCT TCTTCTTG 48

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 48 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GCTCGTTTTT TTTTCTCTG CAGCCCGGGC TCGAGTCACA GCTTCACC 48

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

ACTGACTGCA GCCCGGGCTC GAGGAAGATC AAAAATGGCT CTTC 44

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GAGTAGTGAA CTTTCATGGAT CCTCGAGGTC TTGAAAACCT AGA 43

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 44 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CAATGTCTTG AGAGATCCCG GGATCCTTAA CAACTAGGAA AAGG 44

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GTAAGACACG ACTTATCGCC ACTG 24

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GGAATTTCGTC GACAGATCTC TGCAGCTCGA GGGATCCAAG CTT 43

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: synthetic oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GCTTGTTTCGC CATGGATATC TTCTGTATGT TC 32

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 143 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GAT	GCC	AAA	ANG	CCT	CAC	ATG	CCT	CCT	AGA	GAA	GCT	CAT	GTG	CAA	AAG	48
Asp	Ala	Lys	Xaa	Pro	His	MET	Pro	Pro	Arg	Glu	Ala	His	Val	Gln	Lys	
1				5					10					15		
ACC	CAT	TCA	ATK	CCG	CCT	CAA	AAG	ATT	GAG	ATT	TTC	AAA	TCC	TTG	GAG	96
Thr	His	Ser	Xaa	Pro	Pro	Gln	Lys	Ile	Glu	Ile	Phe	Lys	Ser	Leu	Glu	
			20					25					30			
GGT	TGG	GCT	GAG	GAG	AAT	GTC	TTG	GTG	CAT	CTT	AAA	CCT	GTG	GAG	AA	143
Gly	Trp	Ala	Glu	Glu	Asn	Val	Leu	Val	His	Leu	Lys	Pro	Val	Glu		
		35					40					45				